fpocket: online tools for protein ensemble pocket detection and tracking

Peter Schmidtke[1], Vincent Le Guilloux[2], Julien Maupetit[3] and Pierre Tufféry[3],*

[1]Departament de Fisicoquimica and Institut de Biomedicina (IBUB), Facultat de Farmacia, Universitat de Barcelona, 08028, Barcelona, Spain, [2]ICOA, Institut de chimie organique et analytique, University of Orleans, Orleans and [3]MTi, INSERM UMR-S973 and RPBS, Université Paris Diderot–Paris 7, F75205 Paris, France

ABSTRACT

Computational small-molecule binding site detection has several important applications in the biomedical field. Notable interests are the identification of cavities for structure-based drug discovery or functional annotation of structures. fpocket is a small-molecule pocket detection program, relying on the geometric α-sphere theory. The fpocket web server allows: (i) candidate pocket detection—fpocket; (ii) pocket tracking during molecular dynamics, in order to provide insights into pocket dynamics—mdpocket; and (iii) a transposition of mdpocket to the combined analysis of homologous structures—hpocket. These complementary online tools allow to tackle various questions related to the identification and annotation of functional and allosteric sites, transient pockets and pocket preservation within evolution of structural families. The server and documentation are freely available at http://bioserv.rpbs.univ-paris-diderot.fr/fpocket.

INTRODUCTION

The prediction of functional sites including ligand binding sites or catalytic sites can guide the design of small molecules that could interact with a protein and modulate its function or drive the selection of targeted mutations for protein engineering. It largely relies on the identification and characterization of clefts and cavities in protein structures.

In the past two decades, various approaches have been proposed to the identification of small-molecule binding sites. These encompass geometric analysis of protein surface such as (1–9) see (10) for more references, energy calculations (11,12), the combination of these with information derived from sequences such as residue conservation (13–15), or even meta-methods combining several such approaches to improve binding site prediction (16). Over the last years, however, several new considerations have become of interest. First, the static view of protein pockets is approximative, as is the identification of these in a static image. Differences in the pocket shape and conformation between the apo and holo proteins are known for several proteins, see for example HSP90 (17) or P38 MAP kinase (18). Whether these changes are induced by the ligand or through self-conformational changes is still controversial [e.g. (19,20)]. Last, transient pockets are known to occur on protein surfaces involved in protein–protein interactions (21,22). Current pocket detection approaches provide useful tools for identifying static pockets on static snapshots provided by the Protein Data Bank (PDB). However, very few attempts (21,22), have been made to treat information available in structural families and/or derived from molecular dynamics. Finally, escaping pocket identification from experimental structures alone is also a concern in a context of intensive genome sequencing (23).

Several online services have been proposed for pocket detection such as Q-SiteFinder (11), LIGSITE*csc*(13), CASTp (24), SCREEN (6), PocketDepth (25) or Metapocket (16). These will usually take as input a protein structure and return one or several candidate pockets. In addition to pocket detection, SplitPocket (26) and fPOP (27) provide means of functional inference by comparing the identified patches with those identified over the complete set of PDB structures.

Recently, fpocket, a new program suite allowing pocket detection was introduced (10). The method makes use of Voronoi tesselation and α-spheres to analyse the protein surface. In a reference test set, 94 and 92% of known binding pockets were correctly identified within the best three ranked pockets from the holo and apo proteins. Here, in addition to making fpocket available online, new directions for pocket detection and analysis are proposed.

The first ambition is to analyse pocket dynamics through iterative pocket tracking on a set of PDB snapshots representing various conformational states of the protein of interest. Such an approach allows one to tackle aspects such as pocket flexibility and transientness,

channel opening or ligand-induced conformational changes. The second purpose is to explore cavity conservation among structural families, to identify potential common structural regions of interest. Both aspects are solved using a common grid-based pocket tracking approach over collection of structures.

CONCEPTS AND METHODS

fpocket

fpocket relies on the concept of α-spheres, a concept initiated by Liang and Edelsbrunner (3) and is also used by Chemical Computing Group in the SiteFinder software (http://www.chemcomp.com/). An α-sphere is a sphere in contact with four atoms on its boundary, not containing any internal atom inside. For a protein, very small spheres are located within the protein, whereas large spheres are located at the exterior. Clefts and cavities correspond to spheres of intermediate radii. Thus, it is possible to filter the ensemble of α-spheres defined from the atoms of a protein according to some minimal and maximal radii values in order to address pocket detection. Based on this, we have recently introduced the fpocket package for pocket detection. For more information refer to (10).

Pocket tracking over collection of structural frames

Given a collection of comparable protein structures, such as provided by molecular dynamics or by homology search, one challenge is to track the persistence of pockets within this set of conformations or frames. The approach used can be summarized as an iterative run of fpocket on each frame, followed by a post-analysis using a grid-based approach, as illustrated Figure 1.

In more detail, a 1 Å spaced grid is generated to encompass previously aligned conformers. The grid allows tracking of pockets (α-spheres) in very precise zones over time. On each grid point the α-sphere density of 8Å^3 volume around it is calculated, corresponding to a small box of a 2 Å sized edge. Furthermore, the associated pocket score for each α-sphere near a grid point is tracked following formula (3).

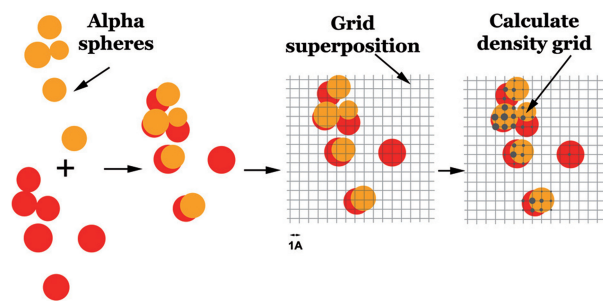Formulas (1) and (3) describe how the densities and scores are calculated.

$$\text{density}(x,y,z) = \frac{1}{n}\sum_{1}^{n} f_{\alpha\text{-spheres}}(x,y,z), \text{ with} \quad (1)$$

$$f_{\alpha\text{-spheres}}(x,y,z) = \text{card}(\alpha\text{-spheres}_{xyz} \\ \cap \{(x,y,z)\pm 1.0\}) \quad (2)$$

$$\text{score}(x,y,z) = \frac{1}{n}\sum_{1}^{n} g_{\alpha\text{-spheres}}(x,y,z), \text{ with} \quad (3)$$

$$g_{\alpha\text{-spheres}}(x,y,z) = \sum \text{pocket}_{\text{score}}(\{\alpha\text{-spheres}_{xyz} \\ \cap \{(x,y,z)\pm 1.0\}\}) \quad (4)$$

(x,y,z) is a given position in the grid space and n is the number of conformations analysed.

Figure 1. Workflow of the pocket tracking methodology. α-spheres from different snapshots are represented by different colors (dark and light).

The resulting grid densities can be used to analyse void space and putative migration pathways of small molecules, whereas scores can rather be used to identify conserved cavities that may bind small molecules. Similar to electron density maps, these pocket density grids can be visualized as iso-volumes, where a given isovalue v allows depiction of all grid points having a density equal to or higher than v.

The interpretation of grid density is often a complex task. Thus, the pocket grid density is mapped to a given reference protein structure using:

$$\text{density}_{\text{score}}(\text{atom}) = \frac{1}{p}\sum_{g_a t_2} \text{score}(\text{grid}_i) \quad (5)$$

where $g_a t_2$ corresponds to the grid points at a distance <2 Å of the atom and p is the number of grid points verifying this condition.

Finally, for visualization purposes using common molecular visualization tools, like PyMol (http://www.pymol.org) or VMD (28), the previously calculated density score for each atom [formula 5] is treated to match a b-factor-like float scale, allowing easy colouring. The expression used to calculate this range of values is:

$$\text{color}(\text{atom}) = \log(1 + \frac{1}{2}\text{density}_{\text{score}}(\text{atom})) \quad (6)$$

mdpocket

mdpocket is the application of the pocket tracking approach to molecular dynamics trajectories. From our experience, we recommend to consider at least 200 snapshots. It can be run in two modes. The first identifies conserved as well as transient pockets and maps them to a pocket density grid. Pocket transientness can be observed by the presence of pockets of lower density in the pocket density grid, whereas high density regions correspond to stable cavities. The second mode requires a user-defined selection of grid points of interest extracted from a previous mdpocket run. It allows to focus on some specific regions of a structure. Tracking of pocket properties on grid selection is performed by considering all α-spheres within the neighbourhood of a selected grid point and merging those into one single pocket. Then all

fpocket descriptors are calculated for this pocket for each frame.

hpocket

hpocket is the application of the pocket tracking approach to collections of homologous structures. Homologous structures are identified either from sequence or from structures. Sequence-based identification is performed using CS-Blast (29) on the PDB (30), filtering the hits in terms of *e*-value, coverage, identity and maximal number. Structure-based identification makes use of the Astral/SCOP (31) classification, using the family level. The hits are then superimposed using an ancillary facility based on TM-align (32).

INPUT/OUTPUT

For all programs of the fpocket suite, two different user interfaces are provided: a classical Common Gateway Interface (CGI) (called *default* interface on the server) and a Mobyle portal (33) interface (called *advanced* interface on the server). In all cases, the same command line is called and generated results are strictly identical. Some advantages of using the Mobyle interface are: (i) the possibility (not mandatory) to open a user session by registering, which allows data persistence on the server; (ii) the possibility to bookmark data for further use, which for mdpocket can avoid the re-upload of possiblity large files; and (iii) results could be directly piped as input to other analysis programs (mdpocket Mode 1 to Mode 2 for example).

fpocket

As input, the fpocket server accepts, a simple standard PDB file or concatenated PDB files to iterate on (each file must start with the HEADER PDB field and ends with the END PDB field). On program termination, for each target, the server returns the results of the stand-alone fpocket program (10), e.g. (i) PyMol and VMD pocket visualization scripts; (ii) the query structure with embedded centers of pocket α-spheres; and (iii) each pocket (set of α-sphere centers) in a PQR file (this modified PDB format allows to set atom van der Waals radius explicitly to determine more precisely the volume detected by fpocket). Additionally, the server provides a set of six snapshots (six sides of a cubic box) showing localization of detected pockets (Figure 2A). Moreover, predicted pockets can be downloaded independently and/or visualized through the embedded Jmol applet (http://jmol.sourceforge.net/) and the OpenAstexViewer (http://www.openastexviewer.net/) for quick analysis of the results (Figure 2B).

mdpocket

The mdpocket server requires a set of PDB snapshots to run the first step of the analysis (Mode 1). At the end of the job, the server proposes three output files: (i) the mdpocket grid that stores density information for each grid point; (ii) the pocket grid points at a particular isovalue (default is 3, i.e. grid points having 3 or more Voronoi vertices in the $8\,\text{Å}^3$ volume around the grid point for each snapshot); and (iii) the pocket α-sphere density stored in the b-factor PDB field of the first snapshot. This last file allows to quickly detect regions of interest. mdpocket allows to run a second step (Mode 2) in order to track descriptors evolution of a user-selected pocket region. To do so, the user should provide selected grid points and the previous set of PDB snapshots. As result of running mdpocket in Mode 2, frame-dependent descriptors are provided as downloadable text file, allowing further treatment using spreadsheet or statistical software like R (34). Finally, mdpocket also provides series of pictures giving an overview of the superposed structure conformational space (see legend of Figure 3A) and pocket α-sphere density (see legend of Figure 3B). Similarily to fpocket, the Jmol applet is embedded in the mdpocket results page for a quick analysis of the pockets dynamics (Figure 3C).

hpocket

hpocket provides results similarly to mdpocket with an additional homology search report, data containing the blast report, PDB hits sequences alignment and superimposed structures.

ONLINE MDPOCKET USE CASE

In the following, we detail a use case utilizing mdpocket and more briefly another more classical use case. Similar analyses could be performed in a homology context using hpocket.

A prototype of mdpocket was used to produce results published in (35). Using the same molecular dynamics for the penta-deoxy FB10L mutant of Type 1 non-symbiotic hemoglobin, Ahb1, from *Arabidopsis thaliana*, 17 snapshots equally spaced in time were submitted to the mdpocket server. Within seconds, the result depicted on the Figure 4A is produced (Mode 1). Comparing results obtained by the server with those published in (35) (Figure 4B), one can notice that the main results can be reproduced even using a very low number of snapshots compared to the number used in (35), 17 versus 800.

For this particular mutant it was shown in (35) that the geminate rebinding of carbon monoxide (CO) is not altered compared to the wild-type. Interestingly, the exit path of CO is closed (Figure 5), which alters geminate rebinding rate. However, the presence of the cavity seen beneath the heme, which is not observed in the wild-type, can explain retention of CO within the structure, without altering geminate rebinding rates as it can be seen for the HE7L mutant. This cavity is regularly disconnected from the *trans*-location pathway of CO, which is furthermore in concordance with a slower geminate rebinding rate compared to the HE7L mutant. Thus, mdpocket allows rapid evaluation of a bunch of snapshots from a Molecular Dynamics (MD) trajectory for conserved pockets or pockets that appear upon time.

As example for the mdpocket Mode 2, Ahb1 is taken again, this time using 267 snapshots picked up during the whole trajectory. Here mdpocket is used to track the volume of the connection between the pocket above the
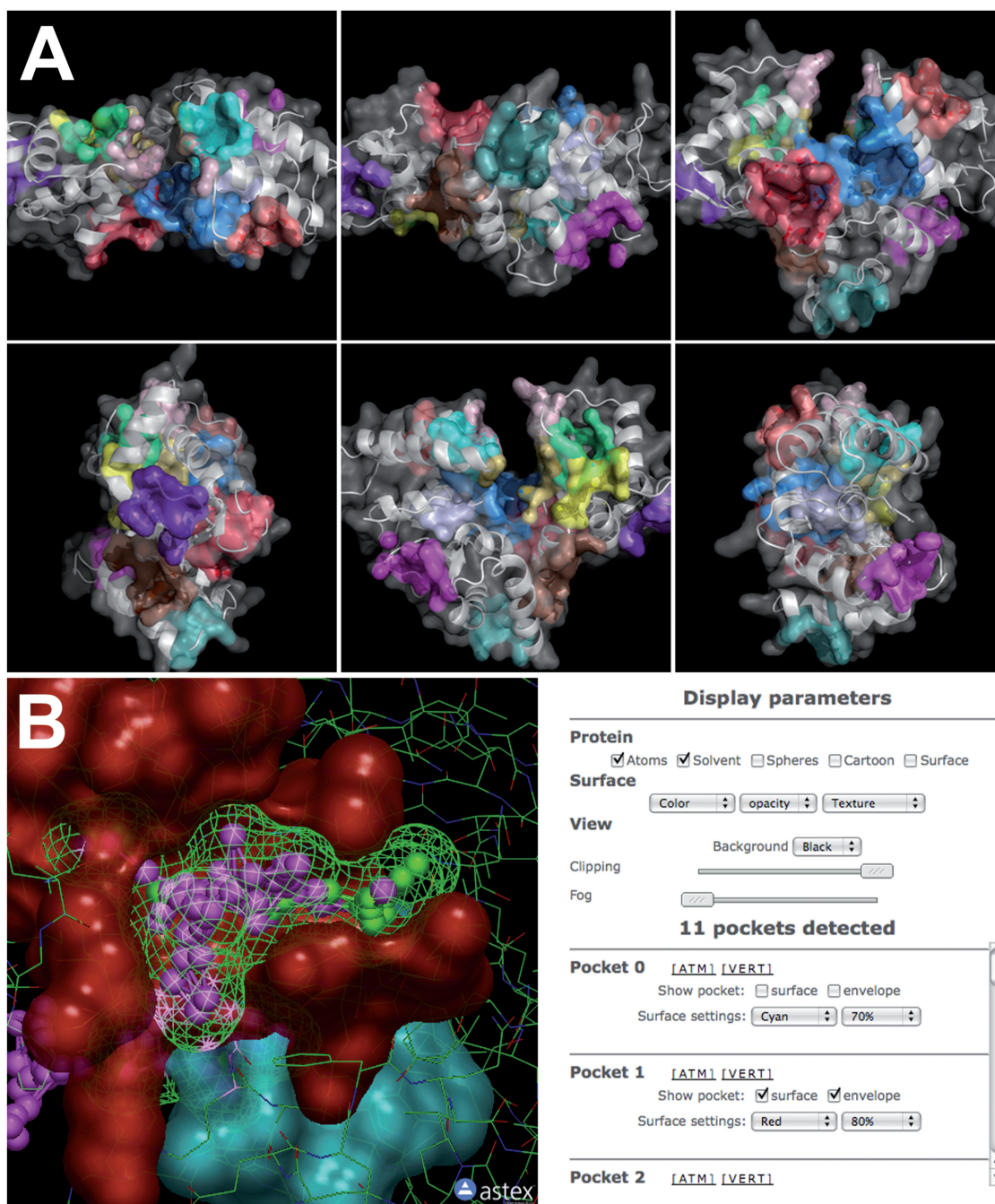
Figure 2. (**A**) fpocket server provides a set of target structure pictures, showing predicted pockets as surfaces of each pockets atoms (one color per pocket). (**B**) The fpocket results page embeds both Jmol and OpenAstexViewer (used here) applets for a quick analysis of the predicted pockets. A control panel on the right part allows the selection of the pockets to visualize and switch between various molecular representations. Here, the surface of the residues of the pocket is in red, the α-spheres in magenta and green and the envelope of the pocket is represented using a green mesh.

heme group and the cavity below the heme group. In order to do so, the selected grid points defining the pocket and the 267 snapshots were uploaded to the mdpocket server. Providing the user-defined pocket definition, mdpocket automatically tracks pocket descriptors over time on this precise zone. The resulting descriptor text file provided by the server was then used to produce the pocket volume curve in Figure 5C. Using the user-defined pocket volume, tracking of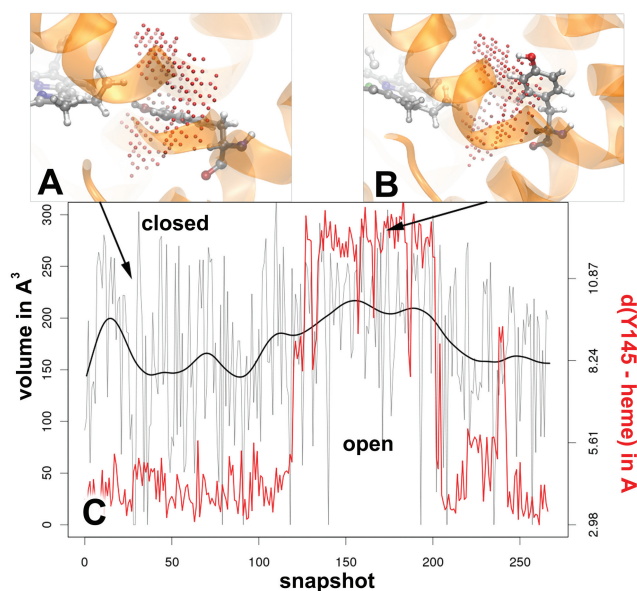 pocket descriptors intends to show transient opening/closing of this channel. A depiction of this zone is available on Figure 5A and B. The heme is situated on the left and the selected pocket grid is represented as red spheres. Furthermore, Y145 is shown in sticks as it plays an interesting role here. This run of mdpocket uses the same input data as the previous, more exploratory phase, just with the additional pocket definition. On Figure 5C, the smoothed volume (black) is plotted over time and despite fluctuations (gray), a mean volume
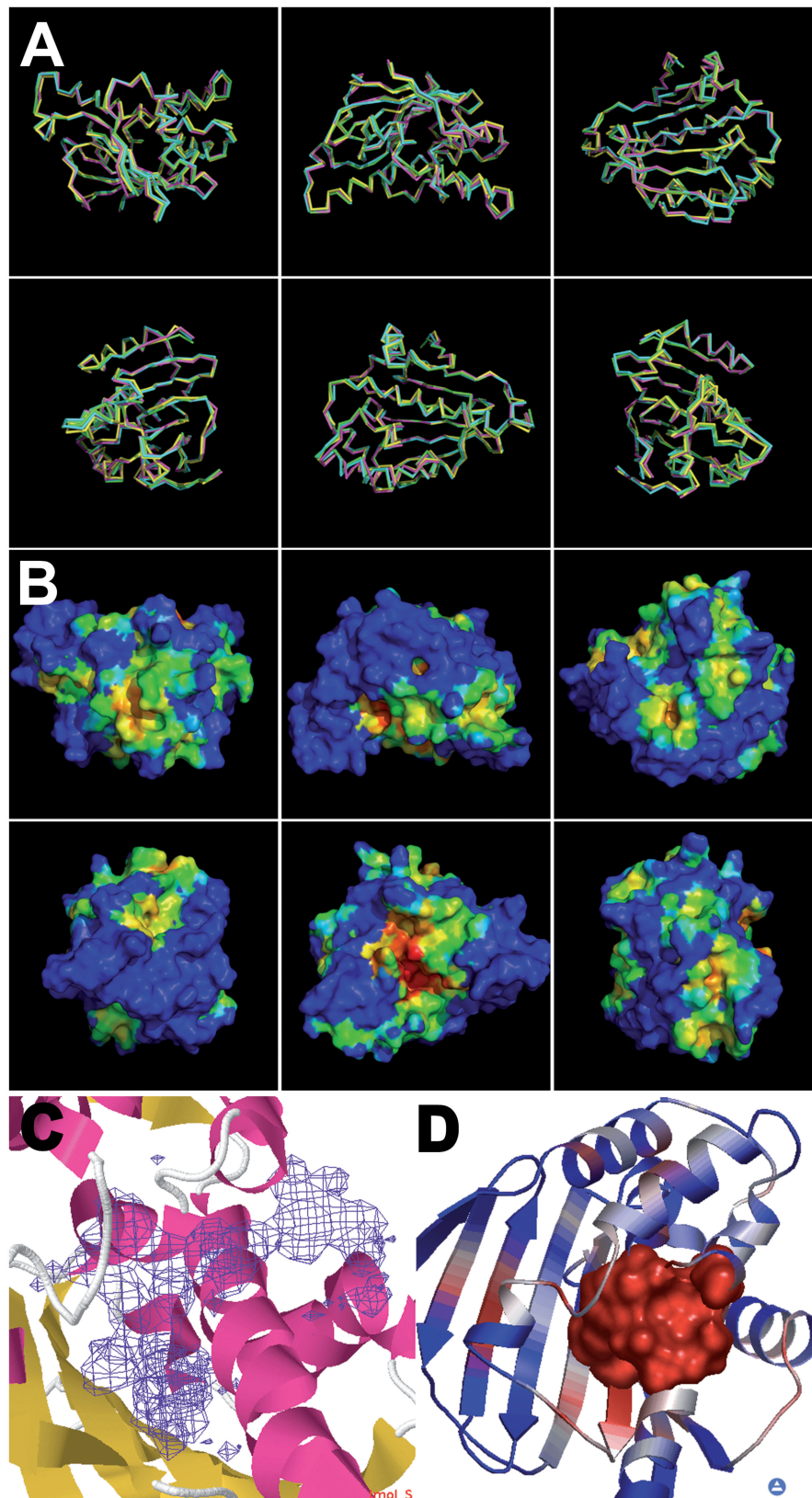
Figure 3. The mdpocket server provides a set of target structure pictures, showing superimposed PDB snapshots as ribbon (**A**) and the first snapshot structure molecular surface colored by α-spheres density ranging from blue (low density) to red (high density) (**B**). The mdpocket results page embeds the Jmol applet (**C**) to give an overview of the conserved cavities (the density grid is represented as an isosurface), and the right part provides viewing components and to extract cavity conservation at a user-selected isovalue—for more details see the main text. It is possible (**D**) to map the density information onto the residues to explore pocket stability in the structure. Here, residues corresponding to high-density regions are displayed as molecular surfaces in the OpenAstexViewer.
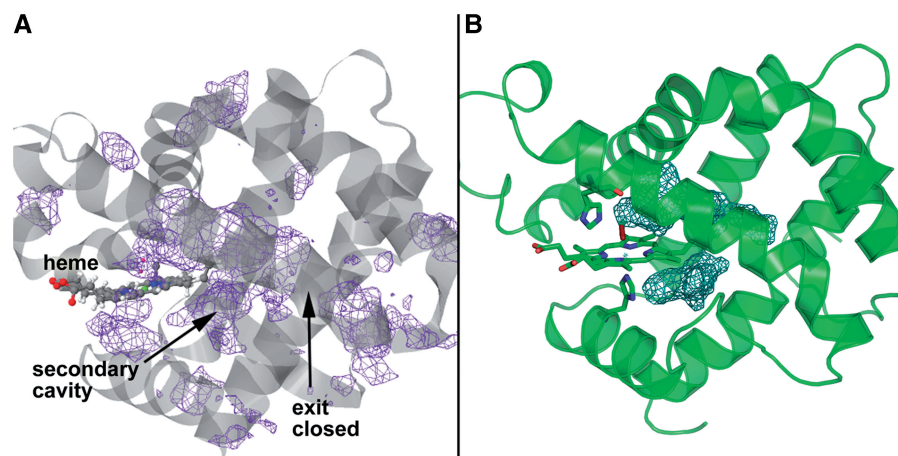
Figure 4. (**A**) Result of a run in Mode 1 of mdpocket on 17 snapshots of a MD trajectory of Ahb1 FB10L. A pocket density grid is provided, allowing visualization of conserved pockets and open channels during the MD trajectory. Analysis shows that the exit pathway of CO is closed in this mutant and that a secondary cavity beneath the heme group is existing. (**B**) Illustration of the results obtained in (35).

confirm the hypothesis that the connection between the two parts of the pocket bordering the heme is sometimes closed. Here, such a closed state can be seen during a long part of the trajectory, preceding a transient opening, which leads again to the closed state.

In the Ahb1 use case, we have illustrated how to investigate very precisely one tiny cavity. Figure 6 on P38 Map kinase shows another use case focusing on the volume of the binding site. Out of a 50 ns trajectory with explicit solvent, 1000 equally spaced snapshots were uploaded to the mdpocket server as well as the previously defined binding site. Several descriptors were tracked for the binding site of P38. Looking at the variation of the binding site volume during the trajectory, an increase of the pocket volume to $1000\,\text{Å}^3$ at around snapshot 200 can be seen. Afterwards, the mean volume of the pocket decreases to $600\,\text{Å}^3$. This very simple example shows how open/closed conformations of cavities can be isolated during MD trajectories, which can have important implications on choosing representative MD snapshots for ensemble docking, for example.

Figure 5. Tracking the volume of the channel between the upper heme pocket and lower heme pocket (mdpocket Mode 2). The pocket grid is shown as red spheres. (**A**) Y145 in the closed state is situated in the pocket, (**B**) Y145 in the open state is situated on the edge pocket. (**C**) smoothed volume of the pocket (black curve) versus time, the distance between the hydroxyl group of Y145 and the heme versus time is represented in red.

increase after 100 snapshots can be noticed. The residue Y145 is situated directly in the selected pocket (Figure 5A) and thus its position towards the pocket is measured using the distance between its hydroxyl group and the proximate heme. One can notice on Figure 5C that the volume increase corresponds to a flipping of Y145 on the side of the cavity (like that seen on figure 5B).

The small volume variation observed despite such important change in torsion angles of Y145 could be explained by the fact that Y145 is still bordering the pocket, but not obstructing it anymore. These results

DISCUSSION AND FUTURE WORK

The fpocket web server provides a valuable, fast, free and easy to use online service allowing to tackle various aspects of protein pocket detection. It relies on a fast and efficient approach for pocket detection from a single protein structure. Furthermore, it allows to investigate new directions to explore and analyse structure ensembles.

Pocket tracking capabilities of mdpocket were shown to coincide with experimental results obtained in (35), thus the mdpocket server provides an easy interface to a new methodology for studying pocket conservation and transientness during molecular dynamics trajectories.

Due to the high scalability of the methodology behind mdpocket, it can also be used to assess pocket conservation among structural families. This functionality is provided by the hpocket service. As suggested by recent pocket detection methods such as Concavity (14) or
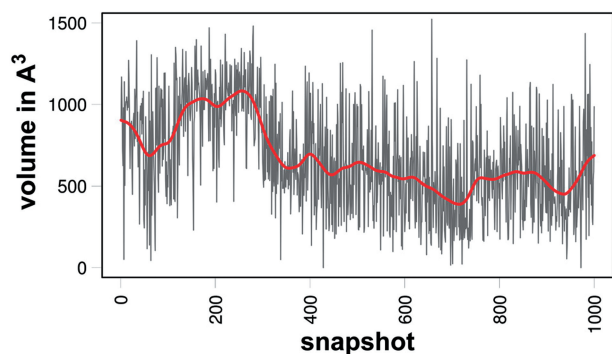
Figure 6. Volume of the P38-binding site during a 50 ns trajectory. Gray line, volume estimated for each snapshot; light gray line: smoothed volume.

Metapocket (16), taking into account the pocket surrounding residue conservation helps to refine precisely a ligand binding site. With hpocket, a complementary approach is proposed, allowing pure geometry-based pocket prediction on homologous structures.

In terms of ergonomics, the design of an online service dealing with such ensembles of structures and analysing dynamic aspects of pockets is a challenging question. The proposed suite of services included in fpocket is, as illustrated in a complex use case considering the analysis of a molecular dynamics trajectory, an efficient starting point. Interestingly, compared to isolated *default* servers, the fpocket, hpocket and mdpocket services are also provided in the Mobyle environment (33). This integration allows easy data pipelining between different applications of the Mobyle portal, including the server presented throughout this article. Presumably, such an integration allows more flexible data handling for the end users, compared to static servers, but could be enhanced. A perspective to gain in interactivity seems to lie in availability of online elementary chainable tools, a direction we are investigating.

REFERENCES

1. Levitt,D.G. and Banaszak,L.J. (1992) Pocket: a computer graphics method for identifying and displaying protein cavities and their surrounding amino acids. *J. Mol. Graph.*, **10**, 229–234.
2. Laskowski,R.A. (1995) SURFNET: a programm for visualizing molecular surfaces, cavities and intermolecular interactions. *J. Mol. Graph.*, **13**, 323–330, 307–308.
3. Liang,J., Edelsbrunner,H. and Woodward,C. (1998) Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein Sci.*, **7**, 1884–1897.
4. Brady,G.P. Jr and Stouten,P.F. (2000) Fast prediction and visualization of protein binding pockets with pass. *J. Comput. Aided Mol. Des.*, **14**, 383–401.
5. Venkatachalam,C.M., Jiang,X., Oldfield,T. and Waldman,M. (2003) Ligandfit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J. Mol. Graph. Model.*, **21**, 289–307.
6. Nayal,M. and Honig,B. (2006) On the nature of cavities on protein surfaces: application to the identification of drug-binding sites. *Proteins*, **6**, 892–906.
7. Weisel,M., Proschak,E. and Schneider,G. (2007) Pocketpicker: analysis of ligand binding-sites with shape descriptors. *Chem. Cent. J.*, **1**, 1–17.
8. Kim,D., Cho,C.-H., Cho,Y., Ryu,J., Bhak,J. and Kim,D.-S. (2008) Pocket extraction on proteins via the voronoi diagram of spheres. *J. Mol. Graph. Model.*, **26**, 1104–1112.
9. Schwarz,B., Albou,L.P., Wurtz,J.M., Pock,O. and Moras,D. (2008) Defining and characterizing protein surface using alpha shapes. *Proteins*, **78**, 1–12.
10. Le Guilloux,V., Schmidtke,P. and Tuffery,P. (2009) Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics.*, **10**, 168.
11. Laurie,A.T.R. and Jackson,R.M. (2005) Q-SteFnder: an energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics.*, **21**, 1908–1916.
12. An,J., Totrov,M. and Abagyan,R. (2005) Pocketome via comprehensive identification and classification of ligand binding envelopes. *Mol. Cell Proteomics*, **4**, 752–761.
13. Huang,B. and Schroeder,M. (2006) LIGSITECSC: predicting ligand binding sites using the connolly surface and degree of conservation. *BMC Struct. Biol.*, **6**, 19.
14. Capra,J.A., Laskowski,R.A., Thornton,J.M., Singh,M. and Funkhouser,T.A. (2009) Predicting protein ligand binding sites by combining evolutionary sequence conservation and 3D structure. *PLoS Comput. Biol.*, **5**, e1000585.
15. Skolnick,J. and Brylinski,M. (2009) Findsite: a combined evolution/structure-based approach to protein function prediction. *Brief. Bioinform.*, **10**, 378–391.
16. Huang,B. (2009) Metapocket: a meta approach to improve protein ligand binding site prediction. *OMICS*, **13**, 325–330.
17. Hessling,M., Richter,K. and Buchner,J. (2009) Dissection of the ATP-induced conformational cycle of the molecular chaperone HSP90. *Nat. Struct. Mol. Biol.*, **16**, 287–293.
18. Sherman,W., Beard,H.S. and Farid,R. (2006) Use of an induced fit receptor structure in virtual screening. *Chem. Biol. Drug Des.*, **67**, 83–84.
19. Brylinski,M. and Skolnick,J. (2008) What is the relationship between the global structures of apo and holo proteins? *Proteins*, **70**, 363–377.
20. Cuneo,M.J., Beese,L.S. and Hellinga,H.W. (2008) Ligand-induced conformational changes in a thermophilic ribose-binding protein. *BMC Struct. Biol.*, **8**, 50.
21. Eyrisch,S. and Helms,V. (2007) Transient pockets on protein surfaces involved in protein-protein interaction. *J. Med. Chem.*, **50**, 3457–3464.
22. Eyrisch,S. and Helms,V. (2009) What induces pocket openings on protein surface patches involved in protein-protein interactions? *J. Comput. Aided Mol. Des.*, **23**, 73–86.
23. Brylinski,M. and Skolnick,J. (2008) A threading-based method (findsite) for ligand-binding site prediction

and functional annotation. *Proc. Natl Acad. Sci. USA*, **105**, 129–134.

24. Dundas,J., Ouyang,Z., Tseng,J., Binkowski,A., Turpaz,Y. and Liang,J. (2006) CASTP: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.*, **34**, W116–W118.

25. Kalidas,Y. and Chandra,N. (2008) PocketDepth: a new depth based algorithm for identification of ligand binding sites in proteins. *J. Struct. Biol.*, **161**, 31–42.

26. Tseng,Y.Y., Dupree,C., Chen,Z.J. and Li,W.-H. (2009) SplitPocket: identification of protein functional surfaces and characterization of their spatial patterns. *Nucleic Acids Res.*, **37**, W384–W389.

27. Tseng,Y.Y., Chen,Z.J. and Li,W.-H. (2010) FPOP: footprinting functional pockets of proteins by comparative spatial patterns. *Nucleic Acids Res.*, **38**, D288–D295.

28. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38, 27–28.

29. Biegert,A. and Söding,J. (2009) Sequence context-specific profiles for homology searching. *Proc. Natl Acad. Sci. USA*, **106**, 3770–3775.

30. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The protein data bank. *Nucleic Acids Res.*, **28**, 235–242.

31. Chandonia,J.-M., Hon,G., Walker,N.S., Lo Conte,L., Koehl,P., Levitt,M. and Brenner,S.E. (2004) The astral compendium in 2004. *Nucleic Acids Res.*, **32**, D189–D192.

32. Zhang,Y. and Skolnick,J. (2005) TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.*, **33**, 2302–2309.

33. Néron,B., Ménager,H., Maufrais,C., Joly,N., Maupetit,J., Letort,S., Carrere,S., Tuffery,P. and Letondal,C. (2009) Mobyle: a new full web bioinformatics framework. *Bioinformatics.*, **25**, 3005–3011.

34. R Development Core Team. (2009) *R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing*. Vienna, Austria, ISBN 3-900051-07-0.

35. Faggiano,S., Abbruzzetti,S., Spyrakis,F., Grandi,E., Viappiani,C., Bruno,S., Mozzarelli,A., Cozzini,P., Astegno,A., Dominici,P. *et al.* (2009) Structural plasticity and functional implications of internal cavities in distal mutants of type 1 non-symbiotic hemoglobin Ahb1 from arabidopsis thaliana. *J. Phys. Chem. B*, **113**, 16028–16038.